Definition 4.1: Matrix

*A **matrix** is a rectangular array of numbers*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

*where the a_{ij} are scalars, called the **entries** or **components** of A. The **size** or **dimension** of a matrix is defined as $m \times n$, where m is the number of rows and n is the number of columns.*

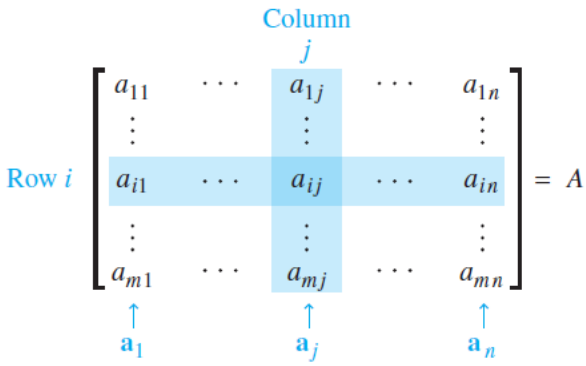


FIGURE 1 Matrix notation.

Definition 4.2: Equality of matrices

*Two matrices are **equal** if they have the same size and the same corresponding entries. More precisely, if $A = [a_{ij}]$ and $B = [b_{ij}]$ are two $m \times n$-matrices, then $A = B$ means that $a_{ij} = b_{ij}$ for all $i = 1, \ldots m$ and $j = 1, \ldots, n$.*

Definition 4.3: Square matrix

*A matrix of size $n \times n$ is called a **square matrix**. In other words, A is a square matrix if it has the same number of rows and columns.*

Definition 4.4: Column vectors and row vectors

*A matrix of size $n \times 1$ is called a **column vector**. A matrix of size $1 \times n$ is called a **row vector**. Here is an example of a column vector X and a row vector Y:*

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}.$$

Definition 4.5: Addition of matrices

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two $m \times n$-matrices. Then $A + B = C$ where C is the $m \times n$-matrix $C = [c_{ij}]$ defined by

$$c_{ij} = a_{ij} + b_{ij}$$

Definition 4.7: The zero matrix

The $m \times n$ ***zero matrix*** *is the $m \times n$-matrix in which all entries are equal to zero. It is denoted by* 0.

Properties

For any matrices A, B and $C \in \mathbb{M}_{m\times n}(\mathbb{K})$

1. $A + B = B + A$. (Commutativity)
2. $A + (B + C) = (A + B) + C$. (Associativity)
3. There exists a matrix $O \in \mathbb{M}_{m\times n}(\mathbb{K})$ with all entries 0 such that $A + O = A$. (Existence of Identity)
4. There exists a matrix $-A$ such that $A + (-A) = O$. (Existence of Inverse)

Remark 1.8 $\mathbb{M}_{m\times n}(\mathbb{K})$ with matrix addition defined on it forms an Abelian group.

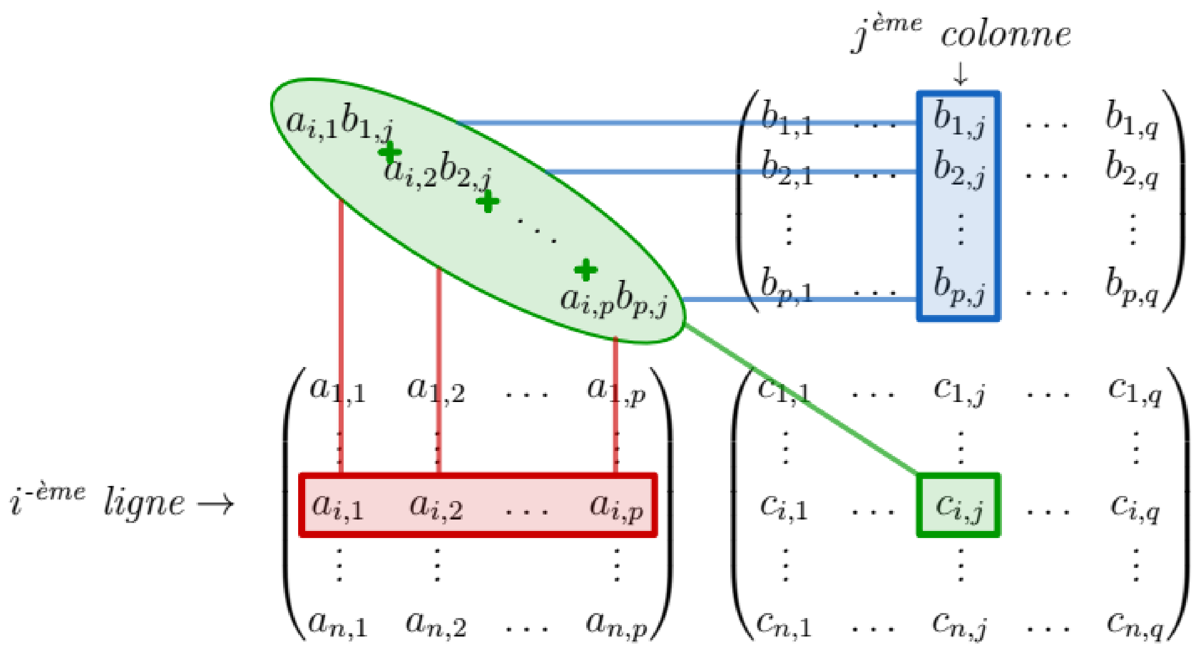
Definition 1.49 (*Matrix Multiplication*) Let $A = \left(a_{ij}\right)_{m\times n}$ and $B = \left(b_{ij}\right)_{n\times p}$. Then their product $AB \in \mathbb{M}_{m\times p}$ and its (i, j)th entry is given by

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

For AB to make sense, the number of columns of A must equal the number of rows of B. Then we say that the size of matrices A and B are compatible for multiplication.

Définition III.1. *Soient $n, p, q \in \mathbb{N}^*$, $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$.*
On définit le ***produit matriciel*** *de A et B, noté $A \times B$ ou AB, comme la matrice $C = (c_{i,j}) \in \mathcal{M}_{n,q}(\mathbb{K})$ définie par :*

$$\forall (i,j) \in [\![1;n]\!] \times [\![1;q]\!],\ c_{i,j} = \sum_{k=1}^{p} a_{i,k} b_{k,j}.$$



Definition 4.12: Scalar multiplication of a matrix

If k is a scalar and $A = \left[a_{ij}\right]$ is a matrix, then $kA = \left[ka_{ij}\right]$.

Proposition 4.14: Properties of scalar multiplication

Let A and B be matrices of the same size, and let k, ℓ be scalars. Then the following properties hold.

- *The distributive law over matrix addition*

$$k(A+B) = kA + kB.$$

- *The distributive law over scalar addition*

$$(k+\ell)A = kA + \ell A.$$

- *The associative law for scalar multiplication*

$$k(\ell A) = (k\ell)A.$$

- *The rule for multiplication by 1*

$$1A = A.$$

Proposition 4.23: Matrix multiplication, column method

Let A be an $m \times n$-matrix, and let B be an $n \times p$-matrix. Suppose that the columns of B are $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p$. Then the columns of AB are

$$A\mathbf{b}_1,\ A\mathbf{b}_2,\ \ldots,\ A\mathbf{b}_p.$$

In other words, the k^{th} column of the matrix product AB is equal to A times the k^{th} column of B.

$$\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 0 \end{bmatrix} = \left[\overbrace{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix}}^{\text{First column}}, \overbrace{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 2 \end{bmatrix}}^{\text{Second column}}, \overbrace{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix}}^{\text{Third column}}, \overbrace{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \end{bmatrix}}^{\text{Fourth column}} \right] = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 4 & 2 & 0 \\ 1 & 2 & 1 & 0 \end{bmatrix}$$

Proposition 4.29: Matrix multiplication, row method

Let A be an $m \times n$-matrix, and let B be an $n \times p$-matrix. Suppose that the rows of A are $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$. Then the rows of AB are

$$\mathbf{a}_1 B,\ \mathbf{a}_2 B,\ \ldots,\ \mathbf{a}_m B.$$

In other words, the i^{th} column of the matrix product AB is equal to the i^{th} column of A times B.

Definition 4.31: Identity matrix

The ***identity matrix*** *of size $n\times n$ has ones along the diagonal, and zeros everywhere else. In other words, it is the matrix $\left[\delta_{ij}\right]$ where $\delta_{ij}=1$ if $i=j$ and $\delta_{ij}=0$ otherwise. The identity matrix is always a square matrix. Here are some identity matrices of various sizes.*

$$\left[\begin{array}{c} 1 \end{array}\right],\quad \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right],\quad \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right],\quad \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

When it is necessary to distinguish which size of identity matrix is being discussed, we will use the notation I_n for the $n\times n$ identity matrix.

Proposition 4.33: Multiplying by the identity matrix

Let A be any $m\times n$-matrix. Then

$$I_mA = A = AI_n.$$

Proposition 4.35: Properties of matrix multiplication

The following properties hold for matrices A,B,C of appropriate dimensions and for scalars r.

- *The associative law of multiplication*

$$(AB)C \;=\; A(BC).$$

- *The existence of multiplicative units*

$$I_mA \;=\; A \;=\; AI_n,$$

  where A is an $m\times n$-matrix.

- *Compatibility with scalar multiplication*

$$(rA)B \;=\; r(AB) \;=\; A(rB).$$

- *The distributive laws of multiplication over addition*

$$\begin{aligned} A(B+C) &= AB+AC, \\ (B+C)A &= BA+CA. \end{aligned}$$

The Transpose of a Matrix

Given an $m\times n$ matrix A, the **transpose** of A is the $n\times m$ matrix, denoted by A^T, whose columns are formed from the corresponding rows of A.

EXAMPLE 8 Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \end{bmatrix}$$

Then

$$A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \quad B^T = \begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix}, \quad C^T = \begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix}$$

Definition 1.51 (*Transpose of a matrix*) The transpose of an $m \times n$ matrix $A = [a_{ij}]$ is the $n \times m$ matrix (denoted by A^T), given by $A^T = [a_{ji}]$.

Properties

Let A and B be matrices of appropriate order, then

1. $(A^T)^T = A$
2. $(A+B)^T = A^T + B^T$
3. $(AB)^T = B^T A^T$
4. $(kA)^T = kA^T$.

Definition 4.68: Symmetric and antisymmetric matrices

An $n \times n$-matrix A is said to be ***symmetric*** *if $A^T = A$. It is said to be* ***antisymmetric*** *(sometimes also called* ***skew symmetric****) if $A^T = -A$.*

Example 4.69: Symmetric and antisymmetric matrices

Let

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 5 & -3 \\ 3 & -3 & 7 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 3 \\ -1 & 0 & 2 \\ -3 & -2 & 0 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 5 & -3 \\ 1 & 3 & 0 \end{bmatrix}.$$

Then A is symmetric because $A^T = A$, B is antisymmetric because $B^T = -B$, and C is neither symmetric nor antisymmetric because C^T is equal to neither C nor $-C$.

Definition 1.52 (*Conjugate transpose of a matrix*) The conjugate transpose of an $m \times n$ matrix $A = [a_{ij}]$ is the $n \times m$ matrix (denoted by A^*) given by $A^* = [\overline{a_{ji}}]$ where bar denotes complex conjugation (if $a_{ij} = c + id$, then $\overline{a_{ij}} = c - id$).

Properties

Let A and B be matrices of appropriate orders and λ be a scalar, then

1. $(A^*)^* = A$
2. $(A+B)^* = A^* + B^*$
3. $(AB)^* = B^* A^*$
4. $(\lambda A)^* = \overline{\lambda} A^*$, where $\overline{\lambda}$ is the conjugate of λ.

Definition 1.53 (*Trace of a matrix*) Let $A = [a_{ij}]$ be an $n \times n$ matrix. The trace of A, denoted by $tr(A)$, is the sum of diagonal entries; that is $tr(A) = \sum_{i=1}^{n} a_{ii}$.

Properties

For any $n \times n$ matrices $A, B, C,$ *and* D and $\lambda \in \mathbb{R}$, we have the following properties:

1. Trace is a linear function.
   $tr(A + B) = tr(A) + tr(B)$
   $tr(\lambda A) = \lambda\, tr(A)$
2. $tr(A^T) = tr(A)$ and $tr(A^*) = \overline{(tr A)}$
3. $tr(AB) = tr(BA)$
4. $tr(ABCD) = tr(DABC) = tr(CDAB) = tr(BCDA)$
5. $tr(ABC) \neq tr(ACB)$ in general.
6. $tr(AB) \neq tr(A).tr(B)$ in general.

Powers of a Matrix

If A is an $n \times n$ matrix and if k is a positive integer, then A^k denotes the product of k copies of A:

$$A^k = \underbrace{A \cdots A}_{k}$$

If A is nonzero and if $\mathbf{x}$ is in $\mathbb{R}^n$, then $A^k\mathbf{x}$ is the result of left-multiplying $\mathbf{x}$ by A repeatedly k times. If $k = 0$, then $A^0\mathbf{x}$ should be $\mathbf{x}$ itself. Thus A^0 is interpreted as the identity matrix. Matrix powers are useful in both theory and applications (Sections 2.6, 4.9, and later in the text).

Definition 4.2 (*Annihilating polynomial*) Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. If for $f(\lambda) \in \mathbb{K}[\lambda]$, we have $f(A) = 0$, then $f(\lambda)$ is called an annihilating polynomial of A.

THE INVERSE OF A MATRIX

An $n \times n$ matrix A is said to be **invertible** if there is an $n \times n$ matrix C such that

$$CA = I \quad \text{and} \quad AC = I$$

where $I = I_n$, the $n \times n$ identity matrix. In this case, C is an **inverse** of A. In fact, C is uniquely determined by A, because if B were another inverse of A, then $B = BI = B(AC) = (BA)C = IC = C$. This unique inverse is denoted by A^{-1}, so that

$$A^{-1}A = I \quad \text{and} \quad AA^{-1} = I$$

A matrix that is *not* invertible is sometimes called a **singular matrix**, and an invertible matrix is called a **nonsingular matrix**.

THEOREM 6

a. If A is an invertible matrix, then A^{-1} is invertible and

$$(A^{-1})^{-1} = A$$

b. If A and B are $n \times n$ invertible matrices, then so is AB, and the inverse of AB is the product of the inverses of A and B in the reverse order. That is,

$$(AB)^{-1} = B^{-1}A^{-1}$$

c. If A is an invertible matrix, then so is A^T, and the inverse of A^T is the transpose of A^{-1}. That is,

$$(A^T)^{-1} = (A^{-1})^T$$

The product of $n \times n$ invertible matrices is invertible, and the inverse is the product of their inverses in the reverse order.

THEOREM 7

An $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n, and in this case, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}.

An Algorithm for Finding A^{-1}

ALGORITHM FOR FINDING A^{-1}

Row reduce the augmented matrix $[\,A \quad I\,]$. If A is row equivalent to I, then $[\,A \quad I\,]$ is row equivalent to $[\,I \quad A^{-1}\,]$. Otherwise, A does not have an inverse.

CHARACTERIZATIONS OF INVERTIBLE MATRICES

THEOREM 8

The Invertible Matrix Theorem

Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.

a. A is an invertible matrix.
b. A is row equivalent to the $n \times n$ identity matrix.
c. A has n pivot positions.
d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
e. The columns of A form a linearly independent set.
f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.
h. The columns of A span $\mathbb{R}^n$.
i. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.
j. There is an $n \times n$ matrix C such that $CA = I$.
k. There is an $n \times n$ matrix D such that $AD = I$.
l. A^T is an invertible matrix.

Let A and B be square matrices. If $AB = I$, then A and B are both invertible, with $B = A^{-1}$ and $A = B^{-1}$.

EXAMPLE 1 Use the Invertible Matrix Theorem to decide if A is invertible:

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix}$$

SOLUTION

$$A \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix}$$

So A has three pivot positions and hence is invertible, by the Invertible Matrix Theorem, statement (c). ■

Multiplication of Partitioned Matrices

Partitioned matrices can be multiplied by the usual row–column rule as if the block entries were scalars, provided that for a product AB, the column partition of A matches the row partition of B.

EXAMPLE 3 Let

$$A = \left[\begin{array}{ccc|cc} 2 & -3 & 1 & 0 & -4 \\ 1 & 5 & -2 & 3 & -1 \\ \hline 0 & -4 & -2 & 7 & -1 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \left[\begin{array}{cc} 6 & 4 \\ -2 & 1 \\ -3 & 7 \\ \hline -1 & 3 \\ 5 & 2 \end{array}\right] = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

The 5 columns of A are partitioned into a set of 3 columns and then a set of 2 columns. The 5 rows of B are partitioned in the same way—into a set of 3 rows and then a set of 2 rows. We say that the partitions of A and B are **conformable** for **block multiplication**. It can be shown that the ordinary product AB can be written as

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix} = \left[\begin{array}{cc} -5 & 4 \\ -6 & 2 \\ \hline 2 & 1 \end{array}\right]$$

$$A_{11}B_1 = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 5 & -2 \end{bmatrix} \begin{bmatrix} 6 & 4 \\ -2 & 1 \\ -3 & 7 \end{bmatrix} = \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix}$$

$$A_{12}B_2 = \begin{bmatrix} 0 & -4 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix}$$

Hence the top block in AB is

$$A_{11}B_1 + A_{12}B_2 = \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix} + \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix} = \begin{bmatrix} -5 & 4 \\ -6 & 2 \end{bmatrix}$$

THEOREM 10

Column–Row Expansion of AB

If A is $m \times n$ and B is $n \times p$, then

$$\begin{aligned} AB &= \begin{bmatrix} \text{col}_1(A) & \text{col}_2(A) & \cdots & \text{col}_n(A) \end{bmatrix} \begin{bmatrix} \text{row}_1(B) \\ \text{row}_2(B) \\ \vdots \\ \text{row}_n(B) \end{bmatrix} \\ &= \text{col}_1(A)\,\text{row}_1(B) + \cdots + \text{col}_n(A)\,\text{row}_n(B) \end{aligned} \tag{1}$$

EXAMPLE 4 Let $A = \begin{bmatrix} -3 & 1 & 2 \\ 1 & -4 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}$. Verify that

$$AB = \text{col}_1(A)\,\text{row}_1(B) + \text{col}_2(A)\,\text{row}_2(B) + \text{col}_3(A)\,\text{row}_3(B)$$

SOLUTION Each term above is an *outer product*. (See Exercises 27 and 28 in Section 2.1.) By the row–column rule for computing a matrix product,

$$\begin{aligned} \text{col}_1(A)\,\text{row}_1(B) &= \begin{bmatrix} -3 \\ 1 \end{bmatrix} \begin{bmatrix} a & b \end{bmatrix} = \begin{bmatrix} -3a & -3b \\ a & b \end{bmatrix} \\ \text{col}_2(A)\,\text{row}_2(B) &= \begin{bmatrix} 1 \\ -4 \end{bmatrix} \begin{bmatrix} c & d \end{bmatrix} = \begin{bmatrix} c & d \\ -4c & -4d \end{bmatrix} \\ \text{col}_3(A)\,\text{row}_3(B) &= \begin{bmatrix} 2 \\ 5 \end{bmatrix} \begin{bmatrix} e & f \end{bmatrix} = \begin{bmatrix} 2e & 2f \\ 5e & 5f \end{bmatrix} \end{aligned}$$

Thus

$$\sum_{k=1}^{3} \text{col}_k(A)\,\text{row}_k(B) = \begin{bmatrix} -3a + c + 2e & -3b + d + 2f \\ a - 4c + 5e & b - 4d + 5f \end{bmatrix}$$

Definition 1.57 (*Rank of a matrix*) The rank of a matrix is the order of the highest order sub-matrix having non-zero determinant.

Properties

1. Let A be an $m \times n$ matrix. Then $Rank(A) \leq min\{m, n\}$.
2. Only zero matrix has rank zero.
3. A square matrix $A_{n \times n}$ is invertible if and only if $Rank(A) = n$.
4. *Sylvester's Inequality:* If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then

$$Rank(A) + Rank(B) - n \leq Rank(AB) \leq min\{Rank(A), Rank(B)\}$$

   This result is named after the famous English mathematician *James Joseph Sylvester (1814–1897).*
5. *Frobenius Inequality:* Let A, B, and C be any matrices such that AB, BC, and ABC exists, then

$$Rank(AB) + Rank(BC) \leq Rank(ABC) + Rank(B)$$

   This result is named after the famous German mathematician *Ferdinand Georg Frobenius (1849–1917).*
6. Rank is sub-additive. That is, $Rank(A + B) \leq Rank(A) + Rank(B)$.
7. $Rank(A) = Rank(A^T) = Rank(A^T A)$.
8. $Rank(kA) = Rank(A)$ if $k \neq 0$.

Definition 1.60 (*Elementary Operations*) There are three kinds of elementary matrix operations:

(1) Interchanging two rows (or columns).
(2) Multiplying each element in a row (or column) by a non-zero number.
(3) Multiplying a row (or column) by a non-zero number and adding the result to another row (or column).

When these operations are performed on rows, they are called *elementary row operations;* and when they are performed on columns, they are called *elementary column operations.*

Definition 1.61 (*Equivalent matrices*) Two matrices A and B are said to be row(column) equivalent if there is a sequence of elementary row(column) operations that transforms A into B and is denoted by $A \sim B$.

Definition 1.62 (*Row Echelon form of a matrix*) A matrix is said to be in row echelon form when it satisfies the following conditions:

(a) Each leading entry (the first non-zero entry in a row) is in a column to the right of the leading entry in the previous row.
(b) Rows with all zero elements, if any, are below rows having a non-zero element.

If the matrix also satisfies the condition

(c) The first non-zero element in each row, called the leading entry or pivot, is 1.

Then the matrix is in *reduced row echelon form.*

The rank of a matrix is equal to the number of non-zero rows in its row echelon